CLAIMS

What is claimed is:

1. A computerized method comprising:

producing a set of finite mixture models (FMM)s from a set of parameter values and training data using an Expectation Maximization (EM) process;

calculating a minimum description length (MDL) value for each of the set of FMMs; and

selecting a FMM based on the corresponding MDL values.

2. The method of claim 1, wherein each parameter value is a model-complexity parameter for the EM process.

3. The method of claim 1, wherein selecting comprises selecting the FMM from the set of FMMs corresponding to the MDL having a smallest MDL value.

4. The method of claim 1, further comprising applying the FMM from the set of FMMs to vector quantize a stream of data.

5. The method of claim 4, wherein the FMM of the set of FMMs defines a vector quantization codebook.

6. The method of claim 1, further comprising applying the FMM from the set of FMMs to cluster a stream of data.

7. The method of claim 6, wherein the FMM from the set of FMMs defines a cluster pattern.

8. A machine-readable medium having executable instructions to cause a device to perform a method comprising:

producing a set of finite mixture models (FMM)s from a set of parameter values and training data using an Expectation Maximization (EM) process;

calculating a minimum description length (MDL) value for each of the set of FMMs; and

selecting a FMM based on the corresponding MDL values.

9. The machine-readable medium of claim 8, wherein each parameter value is a model-complexity parameter for the EM process.

10. The machine-readable medium of claim 8, wherein selecting comprises selecting the FMM from the set of FMMs corresponding the MDL having a smallest MDL value.

11. The machine-readable medium of claim 8, further comprising applying the FMM from the set of FMMs to vector-quantize a stream of data.

12. The machine-readable medium of claim 11, wherein the FMM from the set of FMMs defines a vector quantization codebook.

13. The machine-readable medium of claim 8, further comprising applying the FMM from the set of FMMs to cluster a stream of data.

14. The machine-readable medium of claim 13, wherein the FMM from the set of FMMs defines a cluster pattern.

15. A system comprising:

a processor coupled to a memory through a bus; and

a model selection process executed by the processor from the memory to cause the processor to produce a set of finite mixture models (FMM)s from a set of parameter values and training data using an Expectation Maximization (EM) process, to calculate a minimum description length (MDL) value for each of the set of FMMs, and to select a FMM based on the corresponding MDL values.

16. The system of claim 15, wherein each parameter value is a model-complexity parameter for the EM process.

17. The system of claim 15, wherein the model selection process further causes the processor, when selecting, to select the FMM from the set of FMMs corresponding to the MDL having a smallest value.

18. The system of claim 15, wherein the model selection process further causes the processor to apply the FMM from the set of FMMs to vector quantize a stream of data.

19. The system of claim 18, wherein the FMM from the set of FMMs defines a vector quantization codebook.

20. The system of claim 15, wherein the model selection process further causes the processor to apply the FMM from the set of FMMs to cluster a stream of data.

21. The system of claim 20, wherein the FMM from the set of FMMs defines a cluster pattern.

22. An apparatus comprising:

means for producing a set of finite mixture models (FMM)s from a set of parameter values and training data using an Expectation Maximization (EM) process;

means for calculating a minimum description length (MDL) value for each of the set of FMMs; and

means for selecting a FMM based on the corresponding MDL values.

23. The apparatus of claim 22, wherein each parameter value is a model-complexity parameter for the EM process.

24. The apparatus of claim 22, wherein the means for selecting comprises means for selecting the FMM from the set of FMMs corresponding to the MDL having a smallest MDL value.

25. The apparatus of claim 22, further comprising a means for applying the FMM from the set of FMMs to vector-quantize a stream of data.

26. The apparatus of claim 25, wherein the FMM from the set of FMMs defines a vector quantization codebook.

27. The apparatus of claim 22, further comprising a means for applying the FMM from the set of FMMs to cluster a stream of data.

28. The apparatus of claim 27, wherein the FMM from the set of FMMs defines a cluster pattern.